Appln. No. 10/007,615
Amendment dated Dec. 7, 2004
Reply to Office Action of Sep. 7, 2004
Docket No. BOC9-2001-0039 (284)

REMARKS/ARGUMENTS

These remarks are made in response to the Office Action of September 7, 2004 (Office Action). As this response is timely filed within the three-month statutory period, no fee is believed due.

In paragraphs 1-8, the Examiner has rejected claims 1-4, 9-10, 12-13, 15-18, and 23 under 35 U.S.C. § 102(b) as being anticipated by U.S. Patent No. 6,092,044 to Baker, et al. (Baker). In paragraphs 9-13, the Examiner has rejected claims 5-6, 8, 14, 19-20, and 22 under 35 U.S.C. § 103(a) as being unpatentable over Baker in view of U.S. Patent No. 6,363,342 to Shaw, et al. (Shaw). In paragraphs 14-15, the Examiner has rejected claims 7, 11, and 21 under 35 U.S.C. § 103(a) as being unpatentable over Baker in view of U.S. Patent No. 5,850,629 to Holm, et al. (Holm).

Independent Claims 1, 10, and 15 have each been amended so as to further distinguish Applicants' invention. Dependent Claims 5, 6, 19, and 20 have been amended for consistency with the claims from which they depend. The amendments are fully supported in the specification. Particular support for the amendments can be found in the portions of the specification referenced below. No new matter has been added by virtue of these amendments.

It may be helpful to briefly review the features of Applicants' invention before addressing the rejections on the art. The Applicants' invention is directed to a method and system for specifying or composing a pronunciation of a portion of text. In particular, the invention provides a graphical tool through which a user is able to compose or specify a pronunciation for a given portion of text by selecting particular phonemes, or speech sounds, to be included within the pronunciation. Accordingly, the pronunciation can include phonemes, the ordering of the phonemes, as well as other pronunciation parameters relating to the stress and prosody of the text. The resulting pronunciation generated by the invention can be used with a speech recognition system or a text-to-speech system.

7

{WP208588;1}

Appln. No. 10/007,615
Amendment dated Dec. 7, 2004
Reply to Office Action of Sep. 7, 2004
Docket No. BOC9-2001-0039 (284)

Independent Claim 1 has been amended to emphasize certain features of Applicants' invention whereby a pronunciation can be composed by graphically presenting at least one activatable visual identifier corresponding to a set of individual phonemes, and, in response to the selection, generating pronunciation information in accordance with the selected visual identifier. As described in particular at page 5, lines 10-12 of Applicants' specification, the pronunciation information can include at least one phoneme from the set of individual phonemes and at least one other pronunciation parameter in addition to the at least one phoneme; that is, an additional phoneme or a prosodic characteristic. Thus, in addition to phonemes, the pronunciation parameters include the ordering of phonemes and various prosodic characteristics of the words the text for which the pronunciation is composed or specified. (Applicants' specification, p. 5, lines 10-12; see, also, p. 7, line 22 through p. 8, line 5.) The pronunciation can be composed on the basis of the pronunciation information. (See, generally, Applicants' specification, p. 5, line 1, through p. 6, line 2.)

Moreover, as explicitly recited in amended Claim 1, the pronunciation can further be based on additional items. (Applicants' specification, p. 5, lines 16-19; and p. 8, lines 1-5.) These additional items allow a user to fine tune the pronunciation being composed and include audibly playing of portions of the pronunciation as the pronunciation is being composed and/or by being presented with an audible or visual rendering of an exemplary word illustrating a particular phoneme. (See, especially, Applicants' specification, p. 8, lines 1-5.) The composition, as stated in amended Claim 1, can further include compiling the pronunciation information.

Independent Claim 10, as amended, is directed to a pronunciation composition tool. The pronunciation composition tool of Claim 10, as amended, includes a processor for composing a pronunciation based upon pronunciation information, and, so that the user can fine tune the pronunciation, based upon at least one of an audible playing of a

8

portion of the pronunciation as it is being composed and/or an audible or visual rendering of an exemplary word illustrating a particular phoneme.

Independent Claim 15, as amended, provides a computer-readable storage medium on which is stored a computer program that includes machine-executable code sections that include computer instructions for composing a pronunciation. The instructions include instructions for composing a pronunciation based upon pronunciation information, and, for fine tuning the pronunciation, based upon at least one of an audible playing of a portion of the pronunciation as it is being composed and/or an audible or visual rendering of an exemplary word illustrating a particular phoneme.

The Baker Reference

Baker's invention is directed to a method of adding a word to a speech recognition vocabulary. Baker creates a collection of possible phonetic pronunciations from a spelling of the word and uses speech recognition techniques to search a word collection and find the word that best matches an utterance of the word. The collection is created by comparing the spelling to a rules list of letter strings with associated phonemes. The list is searched for a letter string from the spelling of length greater than one letter. The collection is limited to phonetic pronunciations containing phonemes associated with the letter string of length greater than one.

In contrast to Applicants' invention for composing or specifying a pronunciation of a portion of text, Baker performs a mere word-phoneme matching. Moreover, to accomplish the matching, the user is forced to rely on a look up table for matching a typed-in word with a phoneme representation of the word. There is no help provided in an intermediate step in the form of pronunciation information, as with Applicants' invention. This means that much of the work needed to accomplish the desired matching in Baker falls on the shoulders of a user. To successfully perform the matching in Baker, the user needs not only to be facile at typing in words, but, more limiting, must be

9

familiar with the particular set of phonemes that are intended to provide a representation of the typed-in word. This limitation can be particularly problematic if the phonemes are based on a foreign language, especially since no intermediate help in the form of pronunciation information is automatically supplied to guide the user, as with Applicants' invention.

Baker fails to teach or suggest each of the features recited in independent Claims 1, 10, and 15, as amended. For example, Baker fails to teach the composing a pronunciation as is done with Applicants' invention. Accordingly, Baker similarly fails to teach or suggest fine tuning a pronunciation as is also done with Applicants' invention. Baker, more fundamentally, fails to teach or suggest providing pronunciation information, the pronunciation providing a basis for composing the pronunciation, as with Applicants' invention. Baker also fails to provide for fine tuning a pronunciation based upon an audible playing of a portion of the pronunciation and/or an audible or visual rendering of an exemplary word illustrative of a particular phoneme, as with Applicants' invention.

The Shaw Reference

Shaw similarly fails to teach the composing of a pronunciation based on automatically supplied pronunciation information, as taught by Applicants' invention. Shaw is directed to an editing tool that is used for developing word-pronunciation pairs based on a spelled word input. Shaw's editing tool includes a transcription generator that receives the spelled word input from the user and generates a list of suggested phonetic transcriptions. The editor displays the list of suggested phonetic transcriptions to the user and provides a mechanism for selecting a pronunciation from the list of suggested phonetic transcriptions.

Shaw, singly and in combination with Baker, fails to teach or suggest each feature recited in amended independent Claims 1, 10, and 15. For example, Shaw not only fails

10

to teach or suggest providing pronunciation information for composing or specifying a pronunciation. Shaw also fails to provide pronunciation information that specifically includes at least one phoneme from a plurality of phonemes and at least one additional pronunciation parameter comprising another phoneme or prosodic characteristic, as explicitly taught by Applicants' invention. Shaw also fails to teach or suggest fine tuning the pronunciation based upon an audible playing of a portion of the pronunciation and/or an audible or visual rendering of an exemplary word that illustrates the sound of a particular phoneme, as also taught by Applicants' invention.

The Holm Reference

Holm is directed to a text-to-speech control system that is intended to allow a user to control text-to-speech conversion functions without prior training. Like both Baker and Shaw, Holm is not directed to specifying or composing a pronunciation of a portion of text.

Holm, singly and in combination with the other references, also fails to teach or suggest each of the features recited in independent Claims 1, 10, and 15, as amended. More particularly, Holm fails to teach or suggest graphically presenting at least one activatable visual identifier corresponding to individual ones of a plurality of phonemes for use in composing a pronunciation, as taught by Applicants' invention and explicitly recited in amended independent Claims 1, 10, and 15. Holm also fails to teach or suggest responding to the selection of a visual identifier comprising at least one phoneme from the plurality of phonemes and at least one additional pronunciation parameter, as also taught by Applicants' invention and recited in each of the independent claims, as amended. Similarly, Holm fails to teach composing a pronunciation based upon pronunciation information and fine tuning the pronunciation based upon at least one of an audible playing of a portion of the pronunciation and an audible or visual rendering of an

11

Appln. No. 10/007,615
Amendment dated Dec. 7, 2004
Reply to Office Action of Sep. 7, 2004
Docket No. BOCS-2001-0039 (284)

exemplary word illustrative of a particular phoneme, as taught by Applicants' invention and recited in each of the amended independent claims.

CONCLUSION

Baker, Shaw, and Holm, individually and in combination, fail to teach or suggest each of the features recited in independent Claims 1, 10, and 15, as amended. It thus follows that the prior art fails to provide a basis for rejecting Claims 1, 10, and 15, as amended, under 35 U.S.C. § 102 or 35 U.S.C. § 103. The prior art also fails to provide a basis for rejecting the dependent claims since each recites yet additional features over those set forth in the independent claims from which each depends. Applicants, therefore, respectfully request that the rejection of each of Claims 1-23 be withdrawn.

Applicants believe that this application is now in full condition for allowance, which action is respectfully requested. Applicants request that the Examiner call the undersigned if clarification is needed on any matter within this Amendment, or if the Examiner believes a telephone interview would expedite the prosecution of the subject application to completion.

Respectfully submitted,

Date: 12/07/04

Gregory A. Nelson, Registration No. 30,577
Richard A. Hinson, Registration No. 47,652
AKERMAN SENTERFITT
Customer No. 40987
Post Office Box 3188
West Palm Beach, FL 33402-3188
Telephone: (561) 653-5000

12

{WP208588;1}